THOUGHTS

ON THE

VALUE AND SIGNIFICANCE

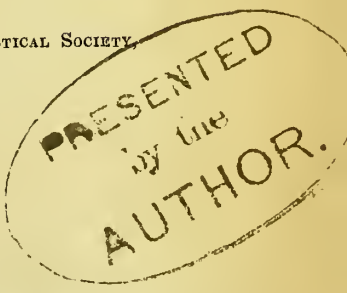
OF

STATISTICS.

By DANIEL NOBLE, M.A., M.D.,

Fellow of the Royal College of Physicians in London.

Read before the Manchester Statistical Society,
January 10th, 1866.



MANCHESTER:
CAVE AND SEVER, PRINTERS BY STEAM POWER, HUNT'S BANK.
1866.

Digitized by the Internet Archive
in 2015

https://archive.org/details/b22350275

THOUGHTS ON STATISTICS.

The value of Statistics is very generally admitted, but their significance is not so commonly understood. That somehow or other they contribute to our useful knowledge, is seen without difficulty; but it is not so readily perceived what they contribute, and in what way. One person, when challenging demonstration, will ask—Can you show that the proposition is supported by statistics? Another, in discussion, will say that statistics can be made to prove anything. Again: whilst some individuals will object to the validity of statistics that they are *no science*, many seem to think that, if you can only show an array of figures on one side, and a certain moral, social, or sanitary condition on the other, you at once make out a causal connection between the numbered facts and the state of affairs brought into conjunction with them. And, indeed, in various ways it is noticeable, how great is the indistinctness of vision when regard is given to statistics.

In this view of things, I have thought that we might pass an agreeable and not unprofitable evening in an attempt to determine the true place of statistics in scientific inquiry—in seeking to ascertain in what manner, and to what extent, the numerical collection of facts may facilitate and lead to useful and practical conclusions.

Our first proceeding in an investigation of this kind, should probably be to define some of the phraseology to be employed. Accordingly, I select the terms *statistics* and *science* for as precise a rendering as I can give, because the terms in question are occasionally used somewhat vaguely, and thus give an ambiguity to discussions of this nature.

I suppose it will be conceded that by statistics we understand an accumulation of facts, classified and enumerated according to their bearings upon some undetermined problem. For example, given the question as to whether or not a particular department of industry is injurious to health and life, to show in figures the extent of the supposed agency side by side with the presumed results, similarly exhibited, in a community subjected to it. A proceeding of this kind is statistical, and facts so numbered constitute statistics.

Now, what is science? I would state it to be an aggregate of facts admitting of generalisation, so as to bring out some law enabling us to predicate their causal relations, and to foresee their reappearance under appreciable circumstances. Thus, Newton's observation of the tendency of bodies to fall, led him so to generalise the facts of gravitation, as to bring out the law according to which the heavenly bodies moved in space. And thus from generalisation of the facts of gravitation, the law was discovered which has rendered astronomy the most perfect of the sciences; a perfection that is attested by the exactness with which deductions can be made from its premises, realisable in accordance with prediction.

Having thus premised, I would say that statistics have value as they are accurate and comparable; and that they have scientific significance as they tend to the solution of some undetermined problem. In more technical phraseology, they may be said to take on the character of science just to the extent that they authorise a generalisation which points to some law.

If these views be just, it is obvious that unless collections of enumerated facts have both accuracy and a certain comparability, they are valueless both for scientific and for practical purposes. And yet, truly, a great deal of information clothed in statistical garb is being constantly published that is quite worthless, either from inaccuracy in the figures, or from particular instances being grouped together that have no proper identity, nor any comparability, when brought to the solution of a difficult question.

The untrustworthiness of some so-called statistics on the score of inaccuracy in the figures, may be familiarly exemplified by what sometimes obtains with our cotton returns; when, for instance, the figures that profess to give the stocks in Liverpool at certain dates are found upon more careful investigation to have been erroneous, and thereby to have misled both the speculator and the more legitimate merchant.

An incongruous grouping of dissimilar facts may be seen in some of the compiled Returns of the causes of death, as they issue from the office of the Registrar-General. Without going to anything like the lengths of our late president, Mr. Aspland, and proclaiming these official statistics to be scientifically valueless, I must yet admit the existence of undoubted shortcomings; these, in the circumstances, would seem to be inevitable; whilst all will acknowledge the earnest care and great judgment which characterise this branch of the public service. However, I proceed to the illustration :—An enumeration of the causes of death under such headings as fever, convulsions, dropsy, and some others, may be regarded as essentially worthless, because the instances so enumerated have no identity rendering them comparable. Fever is a term which designates various and different diseases; dropsy, like cough or headache, is an accidental symptom of maladies having but little in common; and the same thing may be said of convulsions, and of some other terms used in these returns. And if this be a true statement, of what social or sanitary problem can statistics so conditioned assist in the solution? And

if none at all, then these and all others similarly circumstanced can have but little value, and no scientific significance whatever.

But in those cases wherein we can obtain a reliable body of facts numerically arranged with due regard to identity in the separate groupings, and rightly classified with a view to comparison, statistics have great value and much significance. I do not believe that in any instance they are of themselves competent to the definite settlement of a doubtful matter, notwithstanding the frequency of such language as that this or that proposition can be *proved* by statistics. My own conception of their office is, that they are suggestive, auxiliary, and corroborative, but not *probative*. Thus I would say that, when rightly constructed and handled, they may suggest a course of inquiry, or aid in the establishment of a conclusion, or supply corroborative evidence in favour of an inference arrived at by other processes of inquiry; and, again, that they may throw such additional uncertainty upon a proposition already questionable as to exhibit the necessity for other and closer researches.

The history of statistics will furnish many instances of what I now say. A few illustrations of their successful use, and also of their abuse, will both serve to bring out and to exemplify my meaning.

Numbered facts rightly suggested, some years ago, that there was some connection between ill-health and an atmosphere vitiated by the decomposition of animal matters; but they did not prove—as by misinterpretation was for some time thought—that the noxious agency was the existing cause of all fevers. Yet they gave impulse and direction to extended and more intimate investigation; and this produced results, corroborated by other statistics, which have brought us much nearer to the truth of this matter.

Many years ago, it was a widely-entertained opinion that prices were all but *exclusively* governed by the amount of coin and bank notes in circulation; and it was accordingly thought that the Bank of England, by *arbitrary* issues of paper, could almost at its pleasure elevate and depress the price of corn and other commodites.

But the thoughtful mind of the late Mr. Thomas Tooke, applying itself to this question, and examining the facts on all sides, detected the fallacy of this opinion, and showed that a true doctrine of *price* involved many very complex considerations, and was not one to be settled by notions so very simple. He confidently appealed to statistics for corroboration of his own views; and, collecting numerous and well-contrived tables of the bank-issues over long years, and contrasting them with others exhibiting the fluctation in prices at corresponding periods, he obtained such corroboration as perfected the disproof of a prevalent and plausible theory.

Thus we may see that, whether regard be had to affirmation or negation, statistics will suggest direction to research and also supply aid to it; and moreover, may furnish confirmation to deductions obtained by other methods of investigation, and so complete a course of proof.

I have spoken of comparability as a needful quality that should attach to groups of numbered facts, if they are to have scientific value or practical significance. I will now somewhat more explicitly deal with the circumstances and conditions of this requirement.

For statistics, then, to have due comparability, they must not only be faithful and exact, but they must have certain obvious relations; that is to say, the figures produced must have some clear bearing upon a distinct proposition. To enumerate cases of sickness, for example, or the facts of popular education, or the instances of social delinquency, without reference to special circumstances or to some thought on which they presumably shed light, is as though we set up a torch in the desert—illumining no objects—only rendering darkness visible.

Although this estimate of the matter may be plain enough simply enunciated, it is yet the truth that facts and figures are frequently accumulated as if they had some positive value in themselves, apart from any generalised relation to the circumstances of experience embodied in an intelligible formula.

Thus then, first of all, facts enumerated and shown in figures must, if they are to have value, be so carefully and honestly got together as to constitute accurate records; and next, if they are to be rendered significant, particular collections of them must have some common character rendering them pertinent to a definite issue. This I proceed to exemplify.

If I seek to determine the comparative criminality of different places, I must be sure that the numerical results which I contrast have reference to the same thing; and that, too, under circumstances sufficiently similar to warrant a safe conclusion. The crime of murder may supply the illustration. I take the government returns of some foreign country; but before I can rightly compare them with those of our own, I must know in what manner the said returns are made up. I must ascertain whether the same things are always included within the same designation. And if I find that our own country classes beneath the term "murder" the cases only of adult homicide under circumstances of malice aforethought, and that the other comprises in some corresponding class every instance of homicide, including "manslaughter" and what in our returns so often comes out as "concealment of birth," the figures thus brought out are of course incomparable; for any result that such a comparison would show could have no pertinence to the true question at issue.

In like manner, the comparability of figures constituting the statistics of general criminality becomes vitiated by various circumstances,—amongst others, by temporary differences in the particular conditions under which the recorded facts of crime have occurred. For instance, if we were to institute a comparison of the number of published crimes in one country at a season of peace and prosperity with the numbers in another at some period of grave social or political difficulty—at one calculated in its very nature to promote offences against law and order, the results so obtained could not of course authorise any general conclusion.

Moreover, the character and quality of police in different countries should always be ascertained in statistical dealings with topics of this description. The reason is plain enough; the more vigilant and effective the public service in this respect is, the fuller, *cæteris paribus*, will be the records of detected crime.

In further illustration of this question of comparability, I will refer to statistics of illegitimacy as bearing upon the comparative immorality of communities. If we should conclude that, because rural districts ordinarily supply a higher percentage of illegitimate births than towns, therefore the inhabitants of the former were less moral than those of the latter, we should be premature in our inference. Who does not see that an extensive prevalence of the so-called "social evil" in large cities, must largely abate the appreciable issues of immorality, and so diminish the registers of illegitimate births?

But upon this subject one of our vice-presidents, Mr. Greaves, has published much striking though very melancholy information. He has startled us by his sad revelations concerning the frequency both of infanticide and fœticide. And how largely such horrid practices will reduce the returns of extra-nuptial births, it need not be pointed out. Mr. Greaves upon this subject makes statements which, if true, show that Charlotte Winsor is but the type of a class; and that she was far from standing alone in the prosecution of her detestable trade. Hence arise many considerations, which should restrain us from hasty conclusions in dealing with statistics that relate to this repulsive topic.

I will now exemplify the inadequacy of mere statistics to the proof of a class of propositions which, nevertheless, they are often cited with great confidence as making good. I select for the especial illustration an unskilful handling of Sanitary Statistics. We will suppose the mortality-returns of some particular locality to have been collated for two distinct periods; the first we will assume to have been a period at which defective ventilation and cleanliness have been conspicuous;

and the later one to have been a period immediately succeeding the accomplishment of great improvements in these respects. In circumstances of this kind, it may be found that in the first period a high death-rate had been prevalent; and that in the second the mortality had been much reduced. If comparisons of this kind bring out such results, the inference ordinarily drawn is, that the lower death-rate is exclusively attributable to amelioration in the so-called sanitary condition of the district. And yet such an inference very often is a great deal too precipitate. It may have happened, and such a thing sometimes does happen, that the beneficial changes in the physical condition of a neighbourhood have brought to it a better and more provident class of inhabitants, less characterised by thoughtless and early marriages, productive of offspring born but to die. Or it may have been that at the earlier period an exceptionally high mortality had obtained from epidemic or seasonal causes, whilst at the later one these may have ceased to operate; and it is familiar experience, that a reduced death-rate systematically ensues upon the abatement of such intercurrent agencies, quite irrespective of improvements as to ventilation and cleanliness. At any rate, before decisive conclusions can be established by such statistics, these possible sources of fallacy must be looked to and vigilantly eliminated.

But in no department of statistics, probably, are there more sources of fallacy than in Medical Statistics. Cases of disease will sometimes be grouped under one designation that, in regard to intimate nature and curability, have but little in common; and that little very likely accidental rather than essential. Then the comparative sequences of different modes of treatment, as tried upon given numbers, will be shown in statistical tables; the facts dealt with, all the while, being quite devoid of all just comparability. And yet such sequences are often published to the world as scientifically-noted results and consequences.

Before a general auditory like the present, I cannot with propriety engage in any close discussion of this subject, nor yet furnish any

very precise details in illustration; I can make myself understood, however, by citing circumstances hypothetically, which yet may not be altogether imaginary.

Oftentimes, the diagnosis between innocent and malignant tumours in their early stages is very difficult to make. One man will think that a certain hard tumour is cancerous, whilst another will deem it to be something much less serious and quite curable. Well, now, suppose that either from ignorance, or dishonesty, or both, all hard tumours are at once and unscrupulously pronounced to be cancer, whilst in presence of judgment and integrity there is discrimination, and but a very small proportion of the whole is declared to be such. Of course, on this hypothesis the caucer-statistics of the ignorant and unscrupulous charlatan will show an immense per-centage of cures, whilst the scientific and honest practitioner will have nothing of the the kind to display. *Ab uno disce omnes.*

For the right appreciation and the just treatment of statistics, it is, then, clear that something more is required than the mere collection of facts and figures. I think I have shewn that, if we would avoid fallacy, we must not bring together for comparison quantities that are incommensurable. And it will have been apparent from what has preceded, that, without some practical acquaintance with all or most of the facts and circumstances relating to an undetermined problem, the statist cannot successfully deal with his materials. He cannot, without such general knowledge of a subject, point out the relation of the facts to some law that is available for thorough insight; an insight that shall enable him to estimate their causal connections, and recognise the conditions of their occurrence. And so I would say, *en resumé*, that statistics have scientific value and practical significance in the measure in which they bring out the general circumstances under which particular phenomena present themselves.

It has at times been hinted that because in this Society we have not of late been collecting facts and figures so exclusively as was

done in the infancy of its existence, we have deviated from the objects with which it was instituted; that because our later Transactions have been replete with contributions chiefly in the form of essays, we have assumed functions outside and beyond our proper one as a statistical society. I submit, however, that this estimate of our position is incorrect. Our fundamental rules distinctly indicate that our business includes the discussion of all those subjects which concern the advancement and well-being of the community, as well as the simple enumeration of such facts as serve to clear them up. Of course, as a statistical society, we deem it well that in all cases our discussions should have for their basis information that is accurate; and, in many instances, it is needful for securing due accuracy that the facts from which we would draw conclusions should be presented more or less in the arithmetic form. I would yet reiterate, that if by statistics some persons persist in implying the purposeless collection of numbered facts, it is a sense in which we in this Society should not consent to understand the term.

Facts, after all, do but furnish the raw material, as it were, out of which general truths, or principles, are elaborated; and without these, mere knowledge, however accurate in its arithmetic presentation, can have but little worth, either for self-edification or for what is understood by practical usefulness. Indeed, I would say that the want of our time is rather the right interpretation of facts and figures than mere meaningless additions to those we already have. "Although," says Mr. Grove, in his "Correlation of Physical Forces," "we may often err on the side of hasty generalisation, we may equal y err on the side of mere elaborate collection of observations, which, though sometimes leading to a valuable result, yet when cumulated without a connecting link, frequently occasion a costly waste of time, and leave the subject to which they refer in greater obscurity than that in which it was involved in their commencement."

What can be the value of observations unless they supply us with guidance to thought? What is the use of knowledge, except as

transformed by reflection into wisdom? It is a great mistake to suppose, that as you cram the mind with all kinds of information, you necessarily constitute an able man. And yet few mistakes are more common. It has the same prevalence as the kindred misapprehension that children become strong very much in proportion as you load their stomachs with food and drink. The fact is, that as with the physical man that which is taken as food does but strengthen as it becomes one with his very structure,—so with the man mental, knowledge does but invigorate according as it assimilates itself to his higher intellectual nature; a thought which is beautifully expressed by Milton, in his "Paradise Lost:"

> "Knowledge is as food, and needs no less
> Her temperance over appetite, to know
> In measure what the mind may well contain;
> Oppresses else with surfeit, and soon turns
> Wisdom to folly, as nourishment to wind."

There can be no true science of social economy, unless its laws are deduced from a wide induction of reliable facts, which, in many cases, can only have reliability when enumerated in appropriate categories. And so, although statistical science rests substantially upon numbered instances, these in themselves, it should always be remembered, constitute no science. Figures indeed are to scientific statistics very much what knowledge is to wisdom—valuable and significant, when susceptible of due transformation. Thus, in Cowper's "Task:"

> "Knowledge, a rude unprofitable mass,
> The mere materials with which wisdom builds,
> Till smoothed, and squared, and fitted to its place,
> Does but encumber when it seems to enrich."

I must now conclude this paper, and, in doing so, I would state my aim to have been this: To bring out the true function of

statistics in thoughtful inquiry; to indicate the processes whereby they may advantageously aid research, and to point out the abuses to which by neglect of these processes they are liable. I may have but very imperfectly succeeded; but I shall have compensation for the pains which I may have taken in the attempt, if I excite an interest in the line of thought here struck out, so as to influence beneficially the direction which the future labours of this Society may take.

Cave and Sever, Printers, Palatine Buildings, Hunt's Bank, Manchester.